AMENDMENTS TO THE SPECIFICATION

Please amend the paragraph on page 9, lines 15-29, as follows:

In certain ~~embodiment~~ embodiments, the invention provides an approach that combines direct cloning of genes encoding novel or desired bioactivities from environmental samples with a high-throughput screening system designed for the rapid discovery of new molecules, for example, enzymes. The approach is based on the construction of environmental "expression libraries" which can represent the collective genomes of numerous naturally occurring microorganisms archived in cloning vectors that can be propagated in *E. coli* or other suitable host cells. Because the cloned DNA can be initially extracted directly from environmental samples, or from isolates of the environmental samples, the libraries are not limited to the small fraction of prokaryotes that can be grown in pure culture. Additionally, a normalization of the environmental DNA present in these samples ~~could~~ allows a more equal representation of the DNA from all of the species present in a sample. Normalization techniques (described below) can dramatically increase the efficiency of finding interesting genes from minor constituents of the sample that may be under-represented by several orders of magnitude compared to the dominant species in the sample. ~~Normilization~~ Normalization can occur in any of the foregoing embodiments following obtaining nucleic acids from the sample or isolate(s).

D1

Please amend the paragraph on page 10, lines 25-30 to page 11, lines 1-12, as follows:

In some instances it is desirable to identify nucleic acid sequences from a mixed population of organisms, isolates, or enriched populations. In this embodiment, it is not necessary to express gene products. Nucleic acid sequences of interest can be identified or "biopanned" by contacting a clone, device (e.g., a gene chip), filter, or nucleic acid sample with a probe labeled with a detectable molecule. The probe will typically have a sequence that is substantially identical to the nucleic acid sequence of interest, Alternatively, the probe will be a fragment or full length nucleic acid sequence encoding a polypeptide of interest. The ~~prove~~ probe and nucleic acids are incubated under conditions and for such time as to allow the probe and a substantially complementary sequence to hybridize. Hybridization stringency will vary depending on, for example, the length and GC content of the probe. Such factors can be determined empirically (See, for example, Sambrook *et*

D2

al., Molecular Cloning --A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989, and Current Protocols in Molecular Biology, M. Ausubel *et al.*, eds., (Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., most recent Supplement)). Once hybridized the complementary sequence can be PCR amplified, identified by hybridization techniques (e.g., exposing the probe and nucleic acid mixture to a film), or detecting the nucleic acid using a chip.

D2

Please amend the paragraph on page 22, lines 13-20, as follows:

A particularly <u>suitable</u> type of vector for use in the invention contains an f-factor origin <u>of</u> replication. The f-factor (or fertility factor) in *E. coli* is a plasmid which effects high frequency transfer of itself during conjugation and less frequent transfer of the bacterial chromosome itself. In a particular embodiment, cloning vectors referred to as "fosmids" or bacterial artificial chromosome (BAC) vectors are used. These are derived from *E. coli* f-factor, which is able to stably integrate large segments of DNA. When integrated with DNA from a mixed uncultured environmental sample, ~~this makes~~ it <u>is</u> possible to achieve large genomic fragments in the form of a stable environmental gene library.

D3

Please amend the paragraph on page 24, lines 5-10, as follows:

In one embodiment, the library can be screened or sorted to enrich for clones containing a sequence or activity of ~~interested~~ <u>interest</u> based on polynucleotide sequences present in the library or clone. Thus, the invention provides methods and compositions useful in screening organisms for a desired biological activity or biological sequence and to assist in obtaining sequences of interest that can further be used in directed evolution, molecular biology, biotechnology and industrial applications.

D4

Please amend the paragraph on page 30, lines 21-23, as follows:

Lipases/esterases. When screening for an enzyme having lipase or esterase activity, an acylated derivatives of fluorescein ~~in~~ <u>is</u> used. The fluorophore is hydrolyzed from the derivative to generate a signal.

D5

Please amend the paragraph on page 31, lines 3-13, as follows:

D6

Typically, the substrates are able to enter the cell and maintain ~~its~~ their presence within the cell for a period sufficient for analysis to occur (e.g., once the substrate is in the cell it does not "leak" back out before reacting with the enzyme being screened to an ~~extend~~ extent sufficient to produce a detectable response). Retention of the substrate in the cell can be enhanced by a variety of techniques. In one method, the substrate compound is structurally modified by addition of a hydrophobic (e.g., alkyl) tail. In another embodiment, a solvent, such as DMSO or glycerol, can be used to coat the exterior of the cell. Also the substrate can be administered to the cells at reduced temperature, which has been observed to retard leakage of substrates from cells. However, entry of the substrate into the cell is not necessary where, for example, the enzyme or polypeptide is secreted, present in a lysed cellular sample or the like, or where the substrate can act externally to the cell (e.g., an extracellular receptor-ligand complex).

Please amend the paragraph on page 61, lines 4-15, as follows:

D7

End-selection may be applied in combination with any method for performing mutagenesis. Such mutagenesis methods include, but are not limited to, methods described herein (*supra* and *infra*). Such methods include, by way of non-limiting exemplification, any method that may be referred herein or by others in the art by any of the following terms: "saturation mutagenesis", "gene site saturation mutagenesis" or "GSSM", "nucleic acid shuffling", "recombination", "re-assembly", "error-prone PCR", "assembly PCR", "sexual or non-error prone PCR", "crossover PCR", "oligonucleotide primer-directed mutagenesis", "recursive (and/or exponential) ensemble mutagenesis (see Arkin and Youvan, 1992)", "cassette mutagenesis", "*in vivo* mutagenesis", and "*in vitro* mutagenesis". Moreover, end-selection may be performed on molecules produced by any mutagenesis and/or amplification method (see, *e.g.*, Arnold, 1993; Caldwell and Joyce, 1992; Stemmer, 1994) following which method it is desirable to select for (including to screen for the presence of) desirable progeny molecules.

Please amend the paragraph on page 69, lines 18-30, as follows:

D8

Beginning in 1988, single-chain analogues of Fv fragments and their fusion proteins have been reliably generated by antibody engineering methods. The first step generally involves

obtaining the genes encoding VH and VL domains with desired binding properties; these V genes may be isolated from a specific hybridoma cell line, selected from a combinatorial V-gene library, or made by V gene synthesis. The single-chain Fv is formed by connecting the component V genes with an oligonucleotide that encodes an appropriately designed linker peptide, such as (Gly-Gly-Gly-Gly-Ser (~~SEQ ID NO:3~~ SEQ ID NO:2)) or equivalent linker peptide(s). The linker bridges the C-terminus of the first V region and N-terminus of the second, ordered as either VH-linker-VL or VL-linker-VH' In principle, the scfv binding site can faithfully replicate both the affinity and specificity of its parent antibody combining site.

D8

Please amend the paragraph on page 90, lines 18-30, as follows:

Sucrose Gradient (2.2 ml) Size Fractionation. Ligation is stopped by heating the sample to 65 °C for 10 minutes. The sample is gently loaded on a 2.2 ml sucrose gradient and centrifuged in a mini-~~ultracentrifuged~~ ultracentrifuge at 45k rpm at 20 °C for 4 hours (no brake). Fractions are collected by puncturing the bottom of the gradient tube with a 20-gauge needle and allowing the sucrose to flow through the needle. The first 20 drops are collected in a Falcon 2059 tube, and then ten 1-drop fractions (labeled 1-10) are collected. Each drop is about 60 µl in volume. Five µl of each fraction are run on a 0.8% agarose gel to check the size. Fractions 1-4 (about 10-1.5 kb) are pooled and, in a separate tube, fractions 5-7 (about 5-0.5 kb) are pooled. One ml of ice cold ethanol is added to precipitate the DNA and then placed on ice for 10 minutes. The precipitate is pelleted by centrifugation in a microcentrifuge at high speed for 30 minutes. The pellets are washed by resuspending them in 1 ml of 70% ethanol and repelleting them by centrifugation in a microcentrifuge at high speed for 10 minutes, and then dried. Each pellet is then resuspended in 10 µl of TE buffer.

D9

Please amend the paragraph on page 92, line 20, to page 93, line 8, as follows:

Prior to library generation, purified DNA can be normalized. DNA is first fractionated according to the following protocol. A sample composed of genomic DNA is purified on a cesium-chloride gradient. The cesium chloride (Rf = 1.3980) solution is filtered through a 0.2 µm filter and 15 ml is loaded into a 35 ml ~~OptiSeal~~ OPTISEAL tube (Beckman). The DNA is added and thoroughly mixed. Ten micrograms of bis-benzimide (Sigma; Hoechst 33258) is

D10

added and mixed thoroughly. The tube is then ~~filled with the filtered cesium chloride solution and spun~~ centrifuged in a Bti50 rotor in a Beckman L8-70 Ultracentrifuge at 33k rpm for 72 hours. Following centrifugation, a syringe pump and fractionator (Brandel Model 186) are used to drive the gradient through an ISCO UA-5~~UV~~ UV absorbance detector set at 280 nm. Peaks representing DNA from organisms present in an environmental sample are obtained. Eubacterial sequences can be detected by PCR amplification of DNA encoding rRNA from a 10-fold dilution of the *E. coli* peak using the following primers to amplify:

D10

Forward primer: 5'-AGAGTTTGATCCTGGCTCAG-3' (~~SEQ ID NO:4~~ SEQ ID NO:3)

Reverse primer: 5'-GGTTACCTTGTTACGACTT-3' (~~SEQ ID NO:5~~ SEQ ID NO:4)

Recovered DNA is sheared or enzymatically digested to 3-6 kb fragments. Lone-linker primers are ligated and the DNA is size-selected. Size-selected DNA is amplified by PCR, if necessary.

Please amend the paragraph on page 93, line 20 to page 94, line 11, as follows:

Plates of the library prepared as described in Example 1 are used to multiply inoculate a single plate containing 200 µl of LB Amp/Meth, glycerol in each well. This step is performed using the High Density Replicating Tool (HDRT) of the Beckman BIOMEK.RTM. with a 1% bleach, water, isopropanol, air-dry sterilization cycle between each inoculation. The single plate is grown for 2h at 37°C and is then used to inoculate two white 96-well ~~Dynatech~~ DYNATECH microtiter daughter plates containing 250 µl of LB Amp/Meth, glycerol in each well. The original single plate is incubated at 37°C for 18h, then stored at -80°C. The two condensed daughter plates are incubated at 37°C also for 18 h. The condensed daughter plates are then heated at 70°C for 45 min. to kill the cells and inactivate the host *E. coli* enzymes. A stock solution of 5 mg/mL morphourea phenylalanyl-7-amino-4-trifluoromethyl coumarin (MuPheAFC, the "substrate") in DMSO is diluted to 600 µM with 50 mM pH 7.5 Hepes buffer containing 0.6 mg/mL of the detergent dodecyl maltoside. Fifty µl of the 600 µM MuPheAFC solution is added to each of the wells of the white condensed plates with one 100 µl mix cycle using the BIOMEK to yield a final concentration of substrate of about 100 µM. The

D11

fluorescence values are recorded (excitation=400 nm, emission=505 nm) on a plate reading fluorometer immediately after addition of the substrate (t=0). The plate is incubated at 70°C for 100 min, then allowed to cool to ambient temperature for 15 additional minutes. The fluorescence values are recorded again (t=100). The values at t=0 are subtracted from the values at t=100 to determine if an active clone is present.